

Probing Dark Matter Haloes with Satellite Kinematics

Frank C. van den Bosch¹, Peder Norberg¹, H.J. Mo², and Xiaohu Yang² [★]

¹*Department of Physics, Swiss Federal Institute of Technology, ETH Hönggerberg, CH-8093, Zurich, Switzerland*

²*Department of Astronomy, University of Massachusetts, 710 North Pleasant Street, Amherst MA 01003-9305, USA*

ABSTRACT

Using detailed mock galaxy redshift surveys we investigate to what extent the kinematics of large samples of satellite galaxies extracted from flux-limited surveys can be used to constrain halo masses. Unlike previous studies, which focussed only on satellites around relatively isolated host galaxies, we try to recover the average velocity dispersion of satellite galaxies in all haloes, as a function of the luminosity of the host galaxy. We show that previous host-satellite selection criteria yield relatively large fractions of interlopers and with a velocity distribution that, contrary to what has been assumed in the past, differs strongly from uniform. We show that with an iterative, adaptive selection criterion one can obtain large samples of hosts and satellites, with strongly reduced interloper fractions, that allow an accurate measurement of $\sigma_{\text{sat}}(L_{\text{host}})$ over two and a half orders of magnitude in host luminosity. We use the conditional luminosity function (CLF) to make predictions, and show that satellite weighting, which occurs naturally when stacking many host-satellite pairs to increase signal-to-noise, introduces a bias towards higher $\sigma_{\text{sat}}(L_{\text{host}})$ compared to the true, host-averaged mean. A further bias, in the same direction, is introduced when using flux-limited, rather than volume-limited surveys. We apply our adaptive selection criterion to the 2dFGRS and obtain a sample of 12569 satellite galaxies and 8132 host galaxies. We show that the kinematics of these satellite galaxies are in excellent agreement with the predictions based on the CLF, after taking account of the various biases. We thus conclude that there is independent dynamical evidence to support the mass-to-light ratios predicted by the CLF formalism.

Key words: galaxies: halos — galaxies: kinematics and dynamics — galaxies: fundamental parameters — galaxies: structure — dark matter — methods: statistical

1 INTRODUCTION

Ever since the “discovery” of dark matter astronomers have attempted to obtain accurate measurements of the masses of the extended dark matter haloes in which galaxies are thought to reside. A detailed knowledge of halo masses around individual galaxies holds important clues to the physics of galaxy formation, and is an essential ingredient of any successful model that aims at linking the observable Universe (i.e., galaxies) to the bedrock of our theoretical framework (i.e. dark matter).

The main challenge in measuring *total* halo masses is to find a suitable, visible tracer at sufficiently large radii in the halo potential well. Traditionally, starting with the actual discovery of evidence for dark matter by Zwicky (1933, 1937), astronomers have used the kinematics of satellite galaxies. Since the number of detectable satellites in individual systems is generally small, this technique is basically

limited to clusters of galaxies (e.g., Carlberg et al. 1996; Carlberg, Yee & Ellingson 1997b) and the local group (e.g., Little & Tremaine 1987; Lin, Jones & Klemola 1995; Evans & Wilkinson 2000; Evans et al. 2000). However, one can stack the data on many host-satellite pairs to obtain *statistical* estimates of halo masses. Pioneering efforts in this direction were made by Erickson, Gottesman & Hunter (1987), Zaritsky et al. (1993, 1997), and Zaritsky & White (1994). Although these studies were typically limited to samples of less than 100 satellites, they nevertheless sufficed to demonstrate the existence of extended massive dark haloes around (spiral) galaxies.

More recently, large, homogeneous galaxy surveys, such as the Sloan Digital Sky Survey (SDSS; York et al. 2000) and the Two Degree Field Galaxy Redshift Survey (2dFGRS; Colless et al. 2001) have dramatically improved both the quantity and quality of data on (nearby) galaxies, thus allowing the construction of much larger samples of host-satellite pairs (McKay et al. 2002; Prada et al. 2003; Brainerd & Specian 2003). Yet, each of these studies has been

[★] E-mail: vdbosch@phys.ethz.ch

extremely conservative in their selection of “isolated” hosts and “tracer” satellites. For example, McKay et al. (2002), Prada et al. (2003) and Brainerd & Specian (2003) used samples with 1225, 2734, and 2340 satellites, respectively. For comparison, the SDSS and 2dFGRS, from which these samples were selected, contain well in excess of one hundred thousand galaxies. The main reason for being so conservative is to prevent too large numbers of interlopers (i.e., satellites that are not bound to their host galaxies, but only appear associated because of projection effects). In addition, the selection criteria are optimized to only select *isolated* galaxies, with the motivation that the dynamics of binary systems, for example, are more complicated.

In this paper we use a different approach and investigate to what extent the kinematics of satellite galaxies may be used to estimate the mean mass-to-light ratio, *averaged over all possible dark matter haloes*. We make the simple ansatz that satellite galaxies are in virial equilibrium within their dark matter potential well. This is motivated by the finding that dark matter subhaloes, which are likely to be associated with satellite galaxies, have indeed been found to be in a steady-state equilibrium (Diemand, Moore & Stadel 2004). Although we acknowledge that not all systems will be fully virialized, we hypothesize that with a sufficiently large sample of hosts and satellites the assumption of virial equilibrium is sufficiently accurate to describe the mean properties. We use detailed mock galaxy redshift surveys (hereafter MGRSs) to optimize the selection criteria for host and satellite galaxies and show that an iterative, adaptive selection criterion is ideal to limit the number of interlopers while still yielding large numbers of hosts and satellites. Applying our selection criteria to the 2dFGRS yields 8132 hosts with 12569 satellites. We show that the kinematics of these satellite galaxies are in good agreement with predictions based on the conditional luminosity function introduced by Yang, Mo & van den Bosch (2003a) and van den Bosch, Yang & Mo (2003a).

This paper is organized as follows. In Section 2 we first describe, in detail, the construction of our MGRSs. In Section 3 we use these MGRSs to compare three different selection criteria for host and satellite galaxies. We investigate the impact of interlopers and non-central hosts on the satellite kinematics, and show how stacking data from flux-limited surveys results in a systematic overestimate of the true average velocity dispersion of satellite galaxies. We show how the CLF can be used to take these biases into account. In Section 4 we apply our selection criteria to the 2dFGRS and compare the results to analytical estimates and to our MGRS. We summarize our findings in Section 5.

Throughout we assume a flat Λ CDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $h = H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1}) = 0.7$ and with initial density fluctuations described by a scale-invariant power spectrum with normalization $\sigma_8 = 0.9$.

2 MOCK GALAXY REDSHIFT SURVEYS

What is the best way to select host and satellite galaxies from redshift surveys such as the 2dFGRS and the SDSS? How does the flux-limited nature of these surveys impact on the results? How do interlopers bias the mass estimates? In order to address these and other questions we use de-

tailed mock galaxy redshift surveys. These have the advantage that (i) we know exactly the input mass-to-light ratios that we aim to recover, (ii) we can mimic realistic host/selection criteria and investigate the impact of interlopers. In addition, MGRSs allow a detailed investigation of the effect of Malmquist bias, various survey incompleteness effects, boundary effects due to the survey geometry, etc.

To construct MGRSs two ingredients are required; a distribution of dark matter haloes and a description of how galaxies of different luminosity occupy haloes of different mass. For the former we use large numerical simulations (see Section 2.1 below), and for the latter the conditional luminosity function $\Phi(L|M)dL$. The conditional luminosity function (hereafter CLF) was introduced by Yang et al. (2003a) and van den Bosch et al. (2003a) as a statistical tool to link galaxies to their dark matter haloes, and describes the average number of galaxies with luminosity $L \pm dL/2$ that reside in a halo of mass M . As shown in these papers, the CLF is well constrained by the 2dFGRS luminosity function of Madgwick et al. (2002) and the correlation lengths as a function of luminosity obtained by Norberg et al. (2002a). Details about the CLF used in this paper can be found in Appendix A.

2.1 Numerical Simulations

The distribution of dark matter haloes is obtained from a set of large N -body simulations (dark matter only) for a Λ CDM ‘concordance’ cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.7$ and $\sigma_8 = 0.9$. The set consists of a total of six simulations with $N = 512^3$ particles each, and is described in more detail in Jing (2002) and Jing & Suto (2002). All simulations consider boxes with periodic boundary conditions; in two cases $L_{\text{box}} = 100h^{-1} \text{ Mpc}$ while the other four simulations all have $L_{\text{box}} = 300h^{-1} \text{ Mpc}$. Different simulations with the same box size are completely independent realizations and are used to estimate uncertainties due to cosmic variance. The particle masses are $6.2 \times 10^8 h^{-1} \text{ M}_\odot$ and $1.7 \times 10^{10} h^{-1} \text{ M}_\odot$ for the small and large box simulations, respectively. In what follows we refer to simulations with $L_{\text{box}} = 100h^{-1} \text{ Mpc}$ and $L_{\text{box}} = 300h^{-1} \text{ Mpc}$ as L_{100} and L_{300} simulations, respectively.

Dark matter haloes are identified using the standard friends-of-friends algorithm (Davis et al. 1985) with a linking length of 0.2 times the mean inter-particle separation. For each individual simulation we construct a catalogue of haloes with 10 particles or more, for which we store the mass, the position of the most bound particle, and the halo’s mean velocity and velocity dispersion. Haloes that are unbound are removed from the sample. In Yang et al. (2003b) we have shown that the resulting halo mass functions are in excellent agreement with the analytical halo mass function given by Sheth, Mo & Tormen (2001a) and Sheth & Tormen (2002).

2.2 Halo Occupation Numbers

Because of the mass resolution of the simulations and because of the completeness limit of the 2dFGRS we adopt a minimum galaxy luminosity of $L_{\text{min}} = 3 \times 10^7 h^{-2} L_\odot$ throughout. The *mean* number of galaxies with $L \geq L_{\text{min}}$

that resides in a halo of mass M follows from the CLF according to:

$$\langle N \rangle_M = \int_{L_{\min}}^{\infty} \Phi(L|M) dL. \quad (1)$$

In order to Monte-Carlo sample occupation numbers for individual haloes one requires the full probability distribution $P(N|M)$ (with N an integer) of which $\langle N \rangle_M$ gives the mean, i.e.,

$$\langle N \rangle_M = \sum_{N=0}^{\infty} N P(N|M) \quad (2)$$

We use the results of Kravtsov et al. (2003), who has shown that the number of *subhaloes* follows a Poisson distribution. In what follows we differentiate between satellite galaxies, which we associate with these dark matter subhaloes, and central galaxies, which we associate with the host halo (cf. Vale & Ostriker 2004). The total number of galaxies per halo is the sum of N_{cen} , the number of central galaxies which is either one or zero, and N_{sat} , the (unlimited) number of satellite galaxies. We assume that N_{sat} follows a Poissonian distribution and require that $N_{\text{sat}} = 0$ whenever $N_{\text{cen}} = 0$. The halo occupation distribution is thus specified as follows: if $\langle N \rangle_M \leq 1$ then $N_{\text{sat}} = 0$ and N_{cen} is either zero (with probability $P = 1 - \langle N \rangle_M$) or one (with probability $P = \langle N \rangle_M$). If $\langle N \rangle_M > 1$ then $N_{\text{cen}} = 1$ and N_{sat} follows the Poisson distribution

$$P(N_{\text{sat}}|M) = e^{-\mu} \frac{\mu^{N_{\text{sat}}}}{N_{\text{sat}}!}, \quad (3)$$

with $\mu = \langle N_{\text{sat}} \rangle_M = \langle N \rangle_M - 1$. As discussed in Kravtsov et al. (2003) the resulting $P(N|M)$ is significantly sub-Poissonian for haloes with small $\langle N \rangle_M$ (i.e., low mass haloes), but approaches a Poissonian distribution for haloes with large $\langle N \rangle_M$. Such $P(N|M)$ is supported by both semi-analytical models and hydrodynamical simulations of structure formation (Seljak 2000; Benson et al. 2000; Scoccimarro et al. 2001; Berlind et al. 2003), and has been shown to yield correlation functions in better agreement with observations than, for example, for a pure Poissonian $P(N|M)$ (Benson et al. 2000; Berlind & Weinberg 2002; Yang et al. 2003a).

2.3 Assigning galaxies their luminosity and type

Since the CLF only gives the *average* number of galaxies with luminosities in the range $L \pm dL/2$ in a halo of mass M , there are many different ways in which one can assign luminosities to the N_i galaxies of halo i and yet be consistent with the CLF. The simplest approach would be to simply draw N_i luminosities (with $L > L_{\min}$) from $\Phi(L|M)$. We refer to this luminosity sampling as ‘random’. Alternatively, one could use a more constrained approach, and, for instance, always demand that the j^{th} brightest galaxy has a luminosity in the range $[L_j, L_{j-1}]$. Here L_j is defined such that a halo has on average j galaxies with $L > L_j$, i.e.,

$$\int_{L_j}^{\infty} \Phi(L|M) dL = j. \quad (4)$$

We refer to this luminosity sampling as ‘constrained’.

We follow Yang et al. (2003b) and adopt an intermediate approach. Throughout we assume that the central galaxy

is the brightest galaxy in each halo, and we draw its luminosity, L_c , constrained. It therefore has an expectation value of

$$\langle L_c \rangle_M = \int_{L_1}^{\infty} \Phi(L|M) L dL \quad (5)$$

The remaining $N_i - 1$ satellite galaxies are assigned luminosities in the range $L_{\min} < L < L_1$ drawn at random from the distribution function $\Phi(L|M)$.

2.4 Assigning galaxies their phase-space coordinates

Next the mock galaxies need to be assigned a position and velocity within their halo. We assume that each dark matter halo has an NFW density distribution (Navarro, Frenk & White 1997) with virial radius r_{vir} , characteristic scale radius r_s , and concentration $c = r_{\text{vir}}/r_s$. Throughout this paper we compute halo concentrations as a function of halo mass using the relation given by Eke, Navarro & Steinmetz (2001), properly accounting for our definition of halo mass[†]. The ‘central’ (brightest) galaxy in each halo is assumed to be located at the halo centre, which we associate with the position of the most bound particle. Satellite galaxies are assumed to follow a radial number density distribution given by

$$n_{\text{sat}}(r) \propto \left(\frac{r}{\mathcal{R}r_s} \right)^{-\alpha} \left(1 + \frac{r}{\mathcal{R}r_s} \right)^{\alpha-3} \quad (6)$$

(limited to $r \leq r_{\text{vir}}$) with α and \mathcal{R} two free parameters. Unless specifically stated otherwise we adopt $\alpha = \mathcal{R} = 1$ for which the number density distribution of satellite galaxies exactly follows the dark matter mass distribution.

Finally, peculiar velocities are assigned as follows. We assume that the ‘central’ galaxy is located at rest with respect to its halo, and set its peculiar velocity equal to the mean halo velocity. Satellite galaxies are assumed to be in a steady-state equilibrium within the dark matter potential well with an isotropic distribution of velocities with respect to the halo centre. As shown by Diemand et al. (2004) this is a good approximation for dark matter subhaloes, and we assume it also applies to satellite galaxies. One dimensional velocities are drawn from a Gaussian

$$f(v_j) = \frac{1}{\sqrt{2\pi}\sigma_{\text{sat}}(r)} \exp\left(-\frac{v_j^2}{2\sigma_{\text{sat}}^2(r)}\right). \quad (7)$$

with v_j the velocity relative to that of the central galaxy along axis j , and $\sigma_{\text{sat}}(r)$ the local, one-dimensional velocity dispersion obtained from solving the Jeans equation

$$\sigma_{\text{sat}}^2(r) = \frac{1}{n_{\text{sat}}(r)} \int_r^{\infty} n_{\text{sat}}(r') \frac{\partial \Psi}{\partial r}(r') dr' \quad (8)$$

with $\Psi(r)$ the gravitational potential (Binney & Tremaine 1987). Substituting eq. (6) for the spatial number density distribution of satellites yields

$$\sigma_{\text{sat}}^2(r) = \frac{c V_{\text{vir}}^2}{\mathcal{R}^2 \mu_1(c)} \left(\frac{r}{\mathcal{R}r_s} \right)^{\alpha} \left(1 + \frac{r}{\mathcal{R}r_s} \right)^{3-\alpha}$$

[†] Throughout this paper halo masses, denoted by M , are defined as the mass inside the radius R_{180} inside of which the average overdensity is 180.

$$\int_{r/r_s}^{\infty} \frac{\mu_1(x)dx}{(x/\mathcal{R})^{\alpha+2}(1+x/\mathcal{R})^{3-\alpha}} \quad (9)$$

with

$$\mu_\alpha(x) = \int_0^x y^{2-\alpha} (1+y)^{\alpha-3} dy, \quad (10)$$

For $\alpha = \mathcal{R} = 1$ the satellite galaxies follow the same density distribution as the dark matter (i.e., no spatial bias), and (9) reduces to the radial velocity dispersion profile of a spherical NFW potential

$$\sigma_{\text{NFW}}^2(r) = \frac{c V_{\text{vir}}^2}{\mu_1(c)} \left(\frac{r}{r_s} \right) \left(1 + \frac{r}{r_s} \right)^2 \int_{r/r_s}^{\infty} \frac{\mu_1(x)dx}{x^3(1+x)^2} \quad (11)$$

(cf. Klypin et al. 1999).

2.5 Creating Mock Surveys

We aim to construct MGRSs with the same selection criteria and observational biases as in the 2dFGRS, out to a maximum redshift of $z_{\text{max}} = 0.15$. We follow Yang et al. (2003b) and stack $4 \times 4 \times 4$ identical L_{300} boxes (which have periodic boundary conditions), and place the virtual observer in the centre. The central $2 \times 2 \times 2$ boxes, are replaced by a stack of $6 \times 6 \times 6$ L_{100} boxes (see Fig. 11 in Yang et al. 2003b). This stacking geometry circumvents possible incompleteness problems in the mock survey due to insufficient mass resolution of the L_{300} simulations, and easily allows us to reach $z_{\text{max}} = 0.15$ in all directions. We mimic the various observational selection and completeness effects in the 2dFGRS using the following steps:

(i) We define a (α, δ) -coordinate frame with respect to the virtual observer at the centre of the stack of boxes, and remove all galaxies that are not located in the areas equivalent to the NGP and SGP regions of the 2dFGRS.

(ii) For each galaxy we compute the redshift as ‘seen’ by the virtual observer. We take the observational velocity uncertainties into account by adding a random velocity drawn from a Gaussian distribution with dispersion 85 km s^{-1} (Colless et al. 2001), and remove those galaxies with $z > 0.15$.

(iii) For each galaxy we compute the apparent magnitude according to its luminosity and distance, to which we add a rms error of 0.15 mag (Colless et al. 2001; Norberg et al. 2002b). Galaxies are then selected according to the position-dependent magnitude limit, obtained using the apparent magnitude limit masks provided by the 2dFGRS team.

(iv) To take account of the completeness level of the 2dFGRS parent catalogue (Norberg et al. 2002b) we randomly remove 9% of all galaxies.

(v) To take account of the position- and magnitude-dependent completeness of the 2dFGRS, we randomly sample each galaxy using the completeness masks provided by the 2dFGRS team.

Each MGRS thus constructed contains, on average, 144000 galaxies, with a dispersion of ~ 2600 due to cosmic variance. As we show in Section 4 this is in perfect agreement with the 2dFGRS. In addition, we verified that the MGRSs also accurately match the clustering properties (see Yang et al. 2003b), the apparent magnitude distribution and

Table 1. Selection Criteria

SC	R_h	$(\Delta V)_h$	f_h	R_s	$(\Delta V)_s$	f_s
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	2.0	1000	2.0	0.5	1000	4.0
2	2.0	2000	1.0	0.5	2000	1.0
3	$0.8\sigma_{200}$	$1000\sigma_{200}$	1.0	$0.15\sigma_{200}$	2000	1.0

Column (1) indicates the ID of the selection criterion, the parameters of which are listed in Columns (2) to (7) and described in the text. σ_{200} is defined as the velocity dispersion of satellite galaxies around the host galaxy of interest in units of 200 km s^{-1} .

the redshift distribution of the 2dFGRS. Thus, overall our MGRSs are fair representations of the 2dFGRS.

3 METHODOLOGY

3.1 Selection Criteria

A galaxy is considered a potential host galaxy if it is at least f_h times brighter than any other galaxy within a volume specified by $R_p < R_h$ and $|\Delta V| < (\Delta V)_h$. Here R_p is the separation projected on the sky at the distance of the candidate host, and ΔV is the line-of-sight velocity difference. Around each potential host galaxy, satellite galaxies are defined as those galaxies that are at least f_s times fainter than their host and located within a volume with $R_p < R_s$ and $|\Delta V| < (\Delta V)_s$. Host galaxies with zero satellite galaxies are removed from the list of hosts.

In total, the selection of hosts and satellites thus depends on six free parameters: R_h , $(\Delta V)_h$ and f_h to specify the population of host galaxies, and R_s , $(\Delta V)_s$ and f_s to specify the satellite galaxies. These parameters also determine the number of interlopers (defined as a galaxy not physically associated with the halo of the host galaxy) and non-central hosts (defined as a host galaxy that is not the brightest, central galaxy in its own halo). Minimizing the number of interlopers requires sufficiently small R_s and $(\Delta V)_s$. Minimizing the number of non-central hosts requires one to choose R_h , $(\Delta V)_h$ and f_h sufficiently large. Of course, each of these restrictions dramatically reduces the number of both hosts and satellites, making the statistical estimates more and more noisy.

We compare three different selection criteria (hereafter SC), that only differ in their values for the six parameters described above (see Table 1), and use our MGRS to investigate the resulting fractions of interlopers and non-central hosts. Most results are summarized in Table 2 and Fig. 1. SC 1 is identical to that used by McKay et al. (2002) and Brainerd & Specian (2003), who investigated the kinematics of satellite galaxies in the SDSS and 2dFGRS, respectively. The same SC was also used by Prada et al. (2003) for their sample 3. It uses a fairly restrictive set of parameters; host galaxies must be at least two times more luminous ($f_h = 2$) than any other galaxy within a volume specified by $R_h = 2h^{-1} \text{ Mpc}$ and $(\Delta V)_h = 1000 \text{ km s}^{-1}$. Satellite galaxies are selected as those galaxies within $R_s = 0.5h^{-1} \text{ Mpc}$ and $(\Delta V)_s = 1000 \text{ km s}^{-1}$ around each host that are at least four times fainter ($f_s = 4.0$) than the host. Applying these

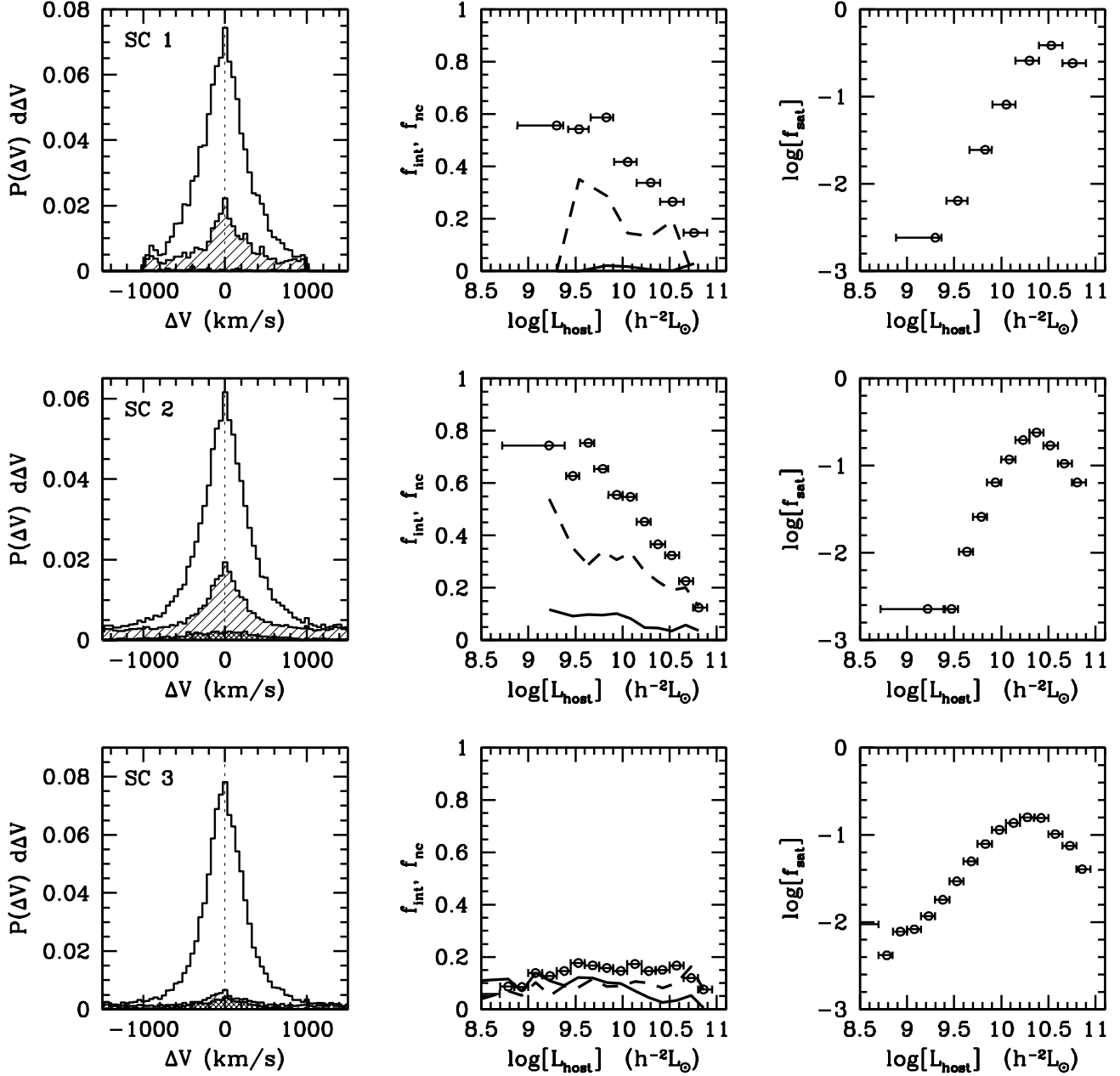


Figure 1. Various statistical properties of host/satellite pairs obtained from our MGRS using three different selection criteria (SC). Panels on the left plot a histogram of the distribution of $\Delta V = V_{\text{sat}} - V_{\text{host}}$. The contributions of interlopers and satellites around non-central hosts are indicated by hatched and cross-hatched histograms, respectively. Note that the $P(\Delta V)$ of interlopers is not uniform, as typically assumed. The open circles in the panels in the middle column indicate the interloper fraction f_{int} for a number of bins in L_{host} . Horizontal errorbars indicate the range of L_{host} used. Dashed lines indicate the interloper fractions obtained by fitting $P(\Delta V)$ with a Gaussian plus a constant (see Section 3.3), which systematically underpredicts the true interloper fraction. Solid lines indicate the fraction of satellites around non-central host galaxies. Panels on the right plot the fraction of satellites in each bin of L_{host} . Note how the number of satellites decreases rapidly with decreasing L_{host} , and that this decrease is less pronounced with SC 3.

SC to our MGRS (which consists of a total of 143727 galaxies), yields 1851 hosts and 3876 satellites. The fraction of interlopers is 27 percent, while only one percent of the hosts is non-central (see Table 2).

The open circles in the upper, middle panel of Fig. 1 show that the interloper fraction, f_{int} , is larger around lower-luminosity hosts, reaching as high as ~ 60 percent around

the faintest hosts in the sample. Clearly, accurate estimates for halo masses based on the kinematics of satellite galaxies requires a proper correction for these interlopers, and thus a detailed knowledge of their velocity distribution, $P_{\text{int}}(\Delta V)$. Thus far, the standard approach has been to assume that $P_{\text{int}}(\Delta V)$ is uniform (i.e., McKay et al. 2002; Brainerd & Specian 2003; Prada et al. 2003). The upper left-hand panel

of Fig. 1 plots $P(\Delta V)$ for host-satellite pairs selected from our MGRS using SC 1. The hatched histogram shows the contribution of interlopers. Clearly, $P_{\text{int}}(\Delta V)$ is not uniform, but instead reveals a pronounced peak around $\Delta V = 0$. This is due to (i) the fact that galaxies, including interlopers, are clustered, and (ii) the infall of galaxies around overdense regions. As we show below, assuming a uniform $P_{\text{int}}(\Delta V)$ in the analysis of satellite kinematics (as is generally done) results in a systematic underestimate of f_{int} but, fortunately, does not lead to significant errors in the kinematics.

With SC 2 our main objective is to increase the number of hosts and satellites, and to include galaxy groups and clusters on a similar footing as ‘isolated’ galaxies. In SC 2 we therefore set both f_h and f_s to unity. This greatly increases the number of both hosts and satellites, and allows brightest cluster galaxies to be included as host galaxies. In order to cover a sufficiently large volume in velocity space to properly sample rich clusters we enlarge both $(\Delta V)_h$ and $(\Delta V)_s$ to 2000 km s^{-1} . As we show below, this has the additional advantage that it allows a better determination of the contribution of interlopers, and therefore a more accurate, statistical correction. SC 2 results in 4 to 5 times as many hosts ($N_{\text{host}} = 7863$) and satellites ($N_{\text{sat}} = 19099$) as with SC 1. However, the fraction of interlopers has also increased, from 27 to 39 percent. Fortunately, as is evident from the middle-left panel of Fig. 1, most of these excess interlopers have $1000 \text{ km s}^{-1} \leq |\Delta V| \leq 2000 \text{ km s}^{-1}$ and are easily corrected for (see Section 3.3). As for SC 1, the fraction of interlopers increases strongly with decreasing L_{host} , with more than 70 percent interlopers around the faintest hosts in the sample. The fraction of non-central hosts, f_{nc} , has increased from one to five percent compared to SC 1. This is mainly a consequence of setting $f_h = 1$, which allows bright galaxies in the outskirts of clusters to be (erroneously) selected as hosts. As evident from the cross-hatched histogram in the middle-left panel of Fig. 1, the $P(\Delta V)$ of satellite galaxies around these non-central hosts is very broad. As for the interlopers, accurate estimates of halo masses based on satellite kinematics require keeping the number of non-central hosts as small as possible.

Both selection criteria discussed so far yield fairly high fractions of interlopers, especially around faint hosts. In addition, as is evident from the panels on the right-hand side of Fig. 1, the number of host-satellite pairs decreases very rapidly for $L_{\text{host}} \lesssim 3 \times 10^{10} h^{-2} L_{\odot}$. This is mainly due to the limited number of faint hosts that make the selection criteria. Ideally, one would use adaptive selection criteria, adjusting $(\Delta V)_h$, R_h and R_s to the virial radius and virial velocity of the halo of the host galaxy in question. This, however, requires prior knowledge of the halo masses as a function of L_{host} , which is exactly what we are trying to recover from the satellite kinematics. We therefore use an iterative procedure: start from an initial guess for $\sigma_{\text{sat}}(L_{\text{host}})$ and estimate the corresponding virial radius and velocity around each individual host galaxy. Use these to adapt $(\Delta V)_h$, R_h , and R_s to the host galaxy in question, and select a new sample of host-satellite pairs. Use the new sample to obtain an improved estimate of $\sigma_{\text{sat}}(L_{\text{host}})$, and start the next iteration. In detail we proceed as follows:

- (i) Use SC 2 to select hosts and satellites

- (ii) Fit the satellite kinematics of the resulting sample with a simple functional form (see Section 3.3).

- (iii) Select new hosts and satellites using $(\Delta V)_h = 1000\sigma_{200} \text{ km s}^{-1}$, $(\Delta V)_s = 2000 \text{ km s}^{-1}$, $R_h = 0.8\sigma_{200}h^{-1} \text{ Mpc}$, and $R_s = 0.15\sigma_{200}h^{-1} \text{ Mpc}$. Here σ_{200} is $\sigma_{\text{sat}}(L_{\text{host}})$ in units of 200 km s^{-1} .

- (iv) Repeat (ii) and (iii) until $\sigma_{\text{sat}}(L_{\text{host}})$ has converged to the required precision. Typically this requires 3 to 4 iterations.

The numerical values in step (iii) are based on extensive tests with our MGRSs, optimizing the results using the following criteria: large N_{sat} and N_{host} , small f_{int} and f_{nc} , and good sampling of L_{host} . The R_h and R_s correspond roughly to 2.0 and 0.375 times the virial radius, respectively. Applying this adaptive SC to our MGRS yields 10483 hosts and 16750 satellites (after 4 iterations). The number of host galaxies has drastically increased with respect to SC 2. As we show below, this allows us to probe the satellite kinematics down to host galaxies with much fainter luminosities. The number of satellite galaxies, on the other hand, has decreased with respect to SC 2. This mainly reflects a drastic decrease in the number of interlopers, from 39 to 15 percent. More importantly, the interloper fraction no longer strongly depends on L_{host} .

3.2 Analytical Estimates

When investigating the impact of interlopers and non-central hosts on the kinematics of satellite galaxies, and comparing different selection criteria, it is useful to have an analytical estimate of the expected $\sigma_{\text{sat}}(L)$. This section describes how the conditional luminosity function (CLF) may be used to compute $\langle \sigma_{\text{sat}}(L) \rangle$ for a flux-limited survey such as the 2dFGRS.

For a halo of mass M , the expectation value for the projected velocity dispersion of satellite galaxies is given by

$$\langle \sigma_{\text{sat}} \rangle_M = \frac{4\pi}{\langle N_{\text{sat}} \rangle_M} \int_0^{r_{\text{vir}}} n_{\text{sat}}(r) \sigma_{\text{sat}}(r) r^2 dr \quad (12)$$

with $\langle N_{\text{sat}} \rangle_M$ the mean number of satellites with $L \geq L_{\text{min}}$ in a halo of mass M which is given by

$$\begin{aligned} \langle N_{\text{sat}} \rangle_M &= \int_{L_{\text{min}}}^{\infty} \Phi(L|M) dL - 1 \\ &\equiv 4\pi \int_0^{r_{\text{vir}}} n_{\text{sat}}(r) r^2 dr \end{aligned} \quad (13)$$

Substituting (6) and (9) yields

$$\langle \sigma_{\text{sat}} \rangle_M = \frac{V_{\text{vir}}}{\mu_{\alpha}(c/\mathcal{R})} \sqrt{\frac{c}{\mathcal{R} \mu_1(c)}} \int_0^{c/\mathcal{R}} \frac{y^{2-\alpha/2} \mathcal{I}^{1/2}(y)}{(1+y)^{(3-\alpha)/2}} dy \quad (14)$$

with

$$\mathcal{I}(y) = \int_y^{\infty} \frac{\mu_1(\mathcal{R}\tau) d\tau}{\tau^{\alpha+2}(1+\tau)^{3-\alpha}} \quad (15)$$

In general, there will not be a one-to-one, purely deterministic, relation between halo mass and host luminosity. Therefore, when averaging over all host galaxies of given luminosity, the expectation value for the velocity dispersion of their satellite galaxies is given by

Table 2. Satellite Kinematics.

Survey (1)	SC (2)	N_{total} (3)	N_{host} (4)	N_{sat} (5)	f_{nc} (6)	f_{int} (7)	$f_{\text{int}}^{\text{fit}}$ (8)	σ_{10} (9)	a_1 (10)	a_2 (11)
MGRS	1	143727	1851	3876	0.01	0.27	0.20	130	0.55	—
MGRS	2	143727	7863	19099	0.05	0.39	0.28	185	0.37	0.17
MGRS	3	143727	10483	16750	0.07	0.15	0.12	176	0.52	0.11
2dFGRS	3	146735	8132	12569	—	—	0.19	193	0.48	0.13

Column (1) indicates whether the sample of host and satellite galaxies has been extracted from our MGRS or from the 2dFGRS. Column (2) indicates the selection criterion used, the parameters of which are listed in Table 1. Columns (3) indicates the total number of galaxies in the survey, while columns (4) and (5) indicate the numbers of host and satellite galaxies, respectively. Columns (6) and (7) indicate the fraction f_{nc} of non-central hosts and the (true) fraction of interlopers f_{int} , both of which are only known for the MGRS. Finally, columns (8) to (11) list the best-fit parameters obtained from the maximum likelihood method described in Section 3.3.

$$\langle \sigma_{\text{sat}}(L_c) \rangle = \int_0^\infty P(M|L_c) \langle \sigma_{\text{sat}} \rangle_M dM \quad (16)$$

with $P(M|L_c)dM$ the conditional probability that a *central* galaxy with luminosity L_c resides in a halo of mass $M \pm dM/2$, and $\langle \sigma_{\text{sat}} \rangle_M$ is the expectation value for the projected velocity dispersion of satellites in a halo of mass M given by eq. (14). Using Bayes theorem, we rewrite (16) as

$$\langle \sigma_{\text{sat}}(L_c) \rangle = \frac{\int_0^\infty P(L_c|M) n(M) \langle \sigma_{\text{sat}} \rangle_M dM}{\int_0^\infty P(L_c|M) n(M) dM} \quad (17)$$

with $n(M)$ the halo mass function and $P(L_c|M)$ the conditional probability that a halo of mass M hosts a central galaxy with luminosity L_c . In our MGRS, L_c is drawn ‘constrained’ for which

$$P(L_c|M) = \begin{cases} \Phi(L_c|M) & \text{if } L_c \geq L_1(M) \\ 0 & \text{if } L_c < L_1(M) \end{cases} \quad (18)$$

with $L_1(M)$ defined by eq. (4). In Appendix B we show how $P(L_c|M)$ can be obtained in the case where L_c is drawn ‘randomly’ (as opposed to ‘constrained’) from the CLF. The dashed curve in Fig 2 plots the host-averaged $\langle \sigma_{\text{sat}}(L_c) \rangle$ thus obtained (see Appendix A for details regarding the CLF used). This $\langle \sigma_{\text{sat}}(L_c) \rangle$ is the true mean velocity dispersion of satellite galaxies around hosts with luminosity L_c , where the mean is taken over the number of host galaxies.

Unfortunately, this is not what an observer who stacks many host-satellite pairs together obtains, which instead is a *satellite-weighted* mean. Since more massive haloes typically contain more satellites, satellite weighting will bias $\langle \sigma_{\text{sat}}(L_c) \rangle$ high with respect to the host-weighted mean. We can use the CLF to estimate the magnitude of this ‘bias’. The satellite-weighted expectation value for $\sigma_{\text{sat}}(L_c)$, in a *volume-limited* survey complete down to a limiting luminosity of L_{min} is

$$\langle \sigma_{\text{sat}}(L_c) \rangle = \frac{\int_0^\infty P(M|L_c) \langle N_{\text{sat}} \rangle_M \langle \sigma_{\text{sat}} \rangle_M dM}{\int_0^\infty P(M|L_c) \langle N_{\text{sat}} \rangle_M dM} \quad (19)$$

The dotted line in Fig. 2 shows the $\langle \sigma_{\text{sat}}(L_c) \rangle$ thus obtained. For $L_{\text{host}} \gtrsim 3 \times 10^9 h^{-2} L_\odot$ the satellite-weighted mean is, as expected, larger than the host-weighted mean (by as much as ~ 40 percent). For less luminous hosts, however, satellite averaging has no significant effect. This is due to the detailed functional form of $\langle N_{\text{sat}} \rangle_M$, which is much shallower at low M than at high M (e.g., van den Bosch et al. 2003a).

The above estimate is based on a complete sampling

of the satellite population of each halo. In reality, however, there are two effects that result in a reduced completeness. First of all, satellites are only selected within a certain projected radius around the host. Whenever that radius is smaller than the virial radius, only a fraction of all satellites enter the sample, and with a mean velocity dispersion that differs from (14). Secondly, in a flux-limited survey the number of satellites around a host of given luminosity depends on redshift. Again, we can account for these two effects using simple algebra.

Suppose we observe a host-satellite system in projection through a circular aperture with radius R_p . The expectation value for the observed velocity dispersion of the satellites in a halo of mass M is given by

$$\langle \sigma_{\text{sat}} \rangle_M = \frac{\int_0^{R_p} dR R \int_R^{r_{\text{vir}}} \frac{r dr}{\sqrt{r^2 - R^2}} n_{\text{sat}}(r) \sigma_{\text{sat}}(r)}{\int_0^{R_p} dR R \int_R^{r_{\text{vir}}} \frac{r dr}{\sqrt{r^2 - R^2}} n_{\text{sat}}(r)} \quad (20)$$

which is straightforward to compute upon substituting (6) and (9). For $R_p = r_{\text{vir}}$ one obtains the total projected velocity given by (14).

The expectation value for a *flux-limited* survey follows from integrating (17) over redshift:

$$\langle \sigma_{\text{sat}}(L_c) \rangle = \frac{1}{V} \int_0^\Omega d\Omega \int_0^{z_{\text{max}}} dz \frac{dV}{d\Omega dz} \langle \sigma_{\text{sat}}(L_c, z) \rangle \quad (21)$$

Here Ω is the solid angle of sky of the survey, dV is the differential volume element, and z_{max} is the minimum of the survey redshift limit (0.15 in our MGRSs) and the maximum redshift out to which a galaxy with luminosity L_c can be detected given the apparent magnitude limit of the survey. The expectation value $\langle \sigma_{\text{sat}}(L_c, z) \rangle$ follows from (17) upon replacing $n(M)$ with $n(M, z)$ (i.e., accounting for the evolution in the halo mass function)[‡]. In addition, since the minimum luminosity of a galaxy in a flux-limited survey depends on redshift the L_{min} in (13) needs to be replaced with $L_{\text{min}}(z)$. Finally, in case $R_p < r_{\text{vir}}$ one needs to use $\langle \sigma_{\text{sat}} \rangle_M$ given by (20) rather than (14).

The solid curve in Fig. 2 shows the $\langle \sigma_{\text{sat}}(L_c) \rangle$ thus obtained with $R_p = r_{\text{vir}}$. Overall, the expectation value for

[‡] Note that we assume here that the CLF does not evolve with redshift, at least not over the small range of redshift ($z \leq 0.15$) considered here.

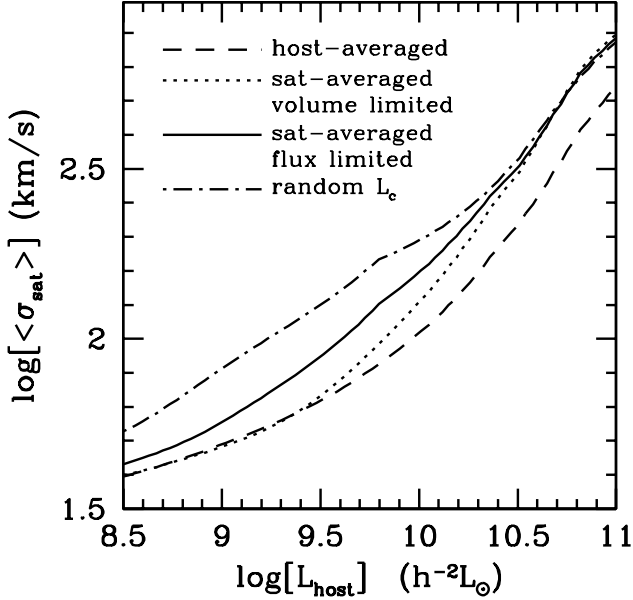


Figure 2. Expectation values for σ_{sat} as a function of L_{host} computed using the CLF as explained in the text. The dashed line corresponds to the true, host-averaged mean. The dotted line shows the satellite-weighted mean, which is what an observer obtains from stacking many host-satellite pairs (using a volume limited survey). The solid and dot-dashed lines show the expectation values for a flux-limited survey, computed using ‘constrained’ and ‘random’ host luminosities, respectively. See text for details.

$\sigma_{\text{sat}}(L_c)$ of a flux-limited survey is larger than for a volume-limited survey. This owes to the fact that, because of the flux limit, smaller mass haloes lose a relatively larger fraction of satellites. Finally, the dot-dashed line shows the same expectation value, satellite-averaged over a flux-limited survey, but computed assuming ‘random’, rather than ‘constrained’, luminosities for the central galaxies. This increases the width of the conditional probability distribution $P(M|L_c)$ (see Appendix B), which in turn strongly increases $\langle\sigma_{\text{sat}}(L_c)\rangle$.

Clearly, satellite weighting from flux-limited surveys introduces large systematic biases in the kinematics of satellite galaxies. The magnitude of this bias depends on, among others, host luminosity and the second moment of $P(M|L_c)$, and can easily be as large as a factor two. Since, to first order, $M \propto \sigma_{\text{sat}}^3$, this implies a systematic overestimate of halo masses of almost an order of magnitude. Clearly, if one were not to correct for these systematic biases, the mass-to-light ratios inferred from satellite kinematics are systematically too high by the same amount. Unfortunately, such bias correction is model-dependent. Although it is straightforward to correct for the bias due to the satellite averaging in a *volume limited survey* (by simply weighting each satellite by the inverse of the number of satellites around the corresponding host galaxy), in a flux-limited survey one has to correct for *missing* satellites (those that did not make the flux-limit). This requires prior knowledge of the abundances and luminosities of satellite galaxies, and is thus model dependent. As shown here, the CLF formalism is one such model that can be used to model these biases in a straightforward way.

3.2.1 The connection between spatial bias and velocity bias

The expectation values discussed so far are based on the assumption that satellite galaxies follow a NFW number density distribution, i.e., we assumed a constant mass-to-number density ratio. Numerous studies have shown this to be a good approximation for clusters of galaxies (e.g., Carlberg et al. 1997a; van der Marel et al. 2000; Lin, Mohr & Stanford 2004; Rhines et al. 2004; Diemand et al. 2004). However, numerical simulations have shown that dark matter subhaloes typically are spatially anti-biased with respect to the mass distribution (i.e., Moore et al. 1998; Ghigna et al. 1999; Colin et al. 1999; Klypin et al. 1999; Okamoto & Habbe 1999; Springel et al. 2001; De Lucia et al. 2004; Diemand et al. 2004). If satellite galaxies follow a similar anti-bias this will reflect itself on the dynamics. We can model spatial anti-bias by adjusting the free parameters α and \mathcal{R} . Setting $\alpha = 0$, for example, introduces a constant number density core, and thus spatial anti-bias at small radii. The parameter \mathcal{R} gives the ratio of the scale radii of satellite galaxies and dark matter, and can be used to adjust the radius of the core region.

Fig. 3 illustrates the effect that spatial bias has on the dynamics of satellite galaxies. We define the *local* velocity bias as

$$b_v(r) \equiv \left[\frac{\sigma_{\text{sat}}^2(r)}{\sigma_{\text{dm}}^2(r)} \right]^{1/2} \quad (22)$$

with $\sigma_{\text{sat}}^2(r)$ and $\sigma_{\text{dm}}^2(r)$ the one-dimensional, isotropic velocity dispersions of satellites and dark matter particles, given by eq. (9) and (11), respectively. If $b_v > 1$ then satellite galaxies typically move faster than dark matter particles in the same halo, and we speak of positive velocity bias. If $b_v < 1$ the satellites are dynamically colder than the dark matter particles, and the velocity bias is said to be negative. In addition to the *local* velocity bias $b_v(r)$, we also define the *global* velocity bias

$$\langle b_v \rangle \equiv \left[\frac{\langle \sigma_{\text{sat}}^2 \rangle}{\langle \sigma_{\text{dm}}^2 \rangle} \right]^{1/2} \quad (23)$$

where $\langle \cdot \rangle$ indicates mass or number-averaged quantities (cf. eq. [12]). The left-hand panel of Fig. 3 plots, for a halo with $M = 10^{12} h^{-1} M_\odot$, the local velocity bias $b_v(r)$ for three different populations of satellite galaxies with $\alpha = 0$ and $\mathcal{R} = 1, 2$, and 3 as indicated. Note the overall positive velocity bias, which reaches very high values at small radii: a (dynamically relaxed) tracer population that is less centrally concentrated than the mass distribution has a positive velocity bias (see also Diemand et al. 2004). The middle panel of Fig. 3 shows that the global velocity bias $\langle b_v \rangle$ is larger for more massive haloes (since these have smaller halo concentrations c). The right-hand panel, finally, plots the expectation values $\langle \sigma_{\text{sat}} \rangle$ as a function of L_{host} . In the case of $\alpha = 0$ and $\mathcal{R} = 3$ the expectation values are a factor 1.1 to 1.2 larger than for the case without spatial bias ($\alpha = \mathcal{R} = 1$). Thus, the effect of spatial (anti)-bias is fairly small; it is much more important to have accurate knowledge of the conditional probability function $P(M|L_c)$ than of $n_{\text{sat}}(r)$ if one is to obtain accurate, unbiased estimates of $\langle \sigma_{\text{sat}} \rangle$ as a function of L_{host} .

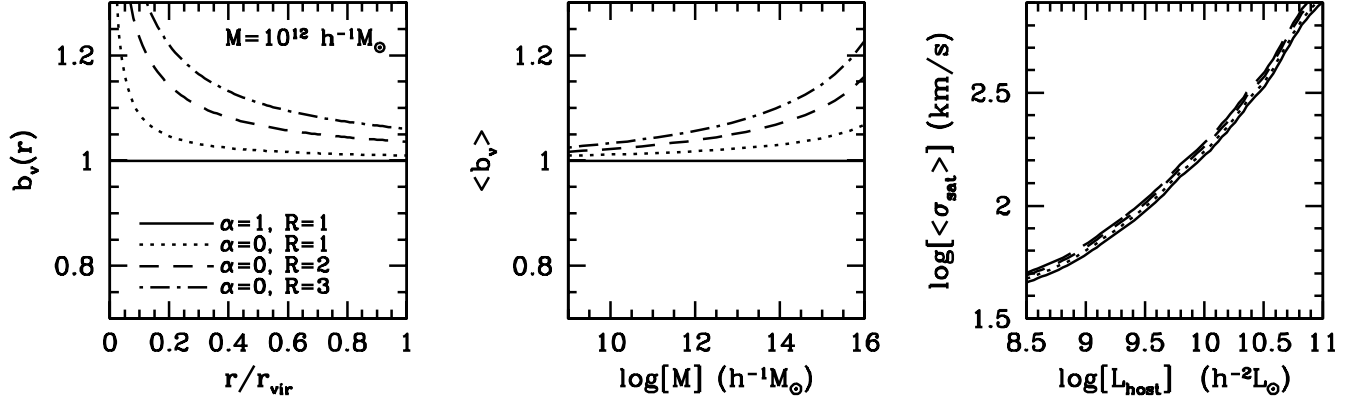


Figure 3. The left-hand panel plots the *local* velocity bias $b_v(r)$ of isotropic, steady-state populations of satellite galaxies that are spatially anti-biased with respect to the mass distribution. Results are shown as a function of radius (normalized to the halo virial radius) for three different spatial distributions, as indicated. For comparison, the solid line corresponds to the case without spatial bias. Note how spatial anti-bias induces a positive velocity bias. The panel in the middle shows the corresponding *global* velocity bias, $\langle b_v \rangle$, as a function of halo mass. More massive haloes reveal a larger, positive velocity bias. Finally, the right-hand panel plots corresponding expectation values for $\sigma_{\text{sat}}(L_{\text{host}})$. The larger the spatial anti-bias, the higher the expectation values, although the effect is marginal.

3.2.2 The impact of orbital anisotropies

So far we have assumed that the orbits of satellite galaxies are isotropic. However, numerical simulations indicate that the orbits of dark matter subhaloes, although close to isotropic near the centre, become slightly radially anisotropic at larger halo-centric radii (e.g. Diemand et al. 2004). In this section we estimate how anisotropy impacts on the projected velocity dispersion of satellite galaxies.

The expectation value for the projected velocity dispersion of satellite galaxies is given by eq. (12). As long as this expectation value integrates over all satellites, it is almost independent of the anisotropy of the orbits. This is most easily seen by considering the virial theorem, which states that for a virialized system $\langle \sigma^2 \rangle_M = |W|/M$, with W the system's total potential energy. Thus, as long as $\langle \sigma \rangle \sim \langle \sigma^2 \rangle^{1/2}$, which is a good approximation for most realistic systems, the projected velocity dispersion of satellite galaxies should be virtually independent of the anisotropy of their orbits.

Contrary to the *global* value of $\langle \sigma^2 \rangle$, the *local* velocity dispersion depends quite strongly on anisotropy. Therefore, as soon as one only considers a radially dependent fraction of the satellites, as is the case with our selection criterion 3 where we integrate over a circular aperture with radius $R_{\text{ap}} \simeq 0.375 R_{\text{vir}}$, the impact of anisotropy is no longer necessarily negligible. In order to estimate the amplitude of this effect we first solve the Jeans equation in spherical symmetry

$$\frac{d}{dr} \rho \sigma_r^2 + \frac{2\beta}{r} \rho \sigma_r^2 + \rho \frac{d\Psi}{dr} = 0 \quad (24)$$

(Binney & Tremaine 1987). Here $\beta(r) = 1 - \frac{\sigma_t^2(r)}{2\sigma_r^2(r)}$ is a measure of the orbital anisotropy, and $\sigma_r(r)$ and $\sigma_t(r)$ are the velocity dispersions in the radial and tangential directions, respectively. Assuming a constant value for β , and using the boundary condition $\rho(r)\sigma^2(r) \rightarrow 0$ for $r \rightarrow \infty$, the solution to this linear differential equation of first order is

$$\rho(r)\sigma_r^2(r) = \frac{G}{r^{2\beta}} \int_r^\infty r'^{2\beta-2} \rho(r') M(r') dr' \quad (25)$$

with G the gravitational constant and $M(r)$ the mass en-

closed within radius r . The expectation value for the projected velocity dispersion integrated over a circular aperture with radius R_{ap} is given by

$$\langle \sigma^2 \rangle(\beta) = \frac{\int_0^{R_{\text{ap}}} dR R \int_R^{R_{\text{vir}}} \rho(r) \sigma_r^2(r) \frac{1-\beta(R/r)^2}{\sqrt{1-(R/r)^2}} dr}{\int_0^{R_{\text{ap}}} dR R \int_R^{R_{\text{vir}}} \frac{\rho(r)}{\sqrt{1-(R/r)^2}} dr} \quad (26)$$

where we have made use of eq [4-60] in Binney & Tremaine (1987). To compute the impact of orbital anisotropy we substitute eq. (25) in the above expression and use an NFW density distribution to compute the ratio

$$\Sigma(\beta) = \left[\frac{\langle \sigma^2(\beta) \rangle}{\langle \sigma^2(0) \rangle} \right]^{1/2} \quad (27)$$

where we set $R_{\text{ap}} = 0.375 R_{\text{vir}}$ as appropriate for our selection criterion 3. We find that Σ increases from 0.97 to 1.05 when β increases from -0.5 to 0.5 . Clearly, even when using an aperture that is only about one-third of the virial radius, orbital anisotropy effects the projected velocity dispersion of satellite galaxies only at the level of a few percent.

3.3 Kinematics

The expectation values discussed above are all based on an idealized situation without interlopers and non-central hosts. We now turn to our MGRSs from which we select hosts and satellites using the selection criteria discussed in Section 3.1. We analyze the kinematics of the satellite galaxies and compare the results to the analytical predictions presented above. This allows us to investigate the impact of interlopers and non-central hosts, and to properly compare the different selection criteria.

The upper panels of Fig 4 show scatter plots of $\Delta V \equiv V_{\text{sat}} - V_{\text{host}}$ as a function of L_{host} , which represents the ‘raw data’ on satellite kinematics that we seek to quantify. We follow McKay et al. (2002) and Brainerd & Specian (2003) and proceed as follows. We bin all host-satellite pairs in a number of bins of L_{host} and fit the distribution $P(\Delta V)$ for each of these bins with the sum of a Gaussian (to represent

the true satellites) plus a constant (to represent the interlopers). The final estimate of the projected velocity dispersion of satellite galaxies, σ_{sat} , follows from the velocity dispersion of the best-fit Gaussian after correcting for the error in ΔV . In our attempt to mimic the 2dFGRS, we added a Gaussian error of 85 km s^{-1} (see Colless et al. 2001) to the velocity of each galaxy in our MGRS. The error on ΔV is therefore equal to $\sqrt{2} \times 85 \text{ km s}^{-1} = 120 \text{ km s}^{-1}$, which we subtract from σ_{sat} in quadrature. The solid circles with errorbars in the lower panels of Fig. 4 plot the $\sigma_{\text{sat}}(L_{\text{host}})$ obtained for the three samples extracted from our MGRS. Vertical errorbars are obtained from the covariance matrix of the Levenberg-Marquardt method used to fit the Gaussian-plus-constant to $P(\Delta V)$, and horizontal errorbars indicate the range of luminosities of the host galaxies in each bin. The dotted, horizontal line indicates the ‘resolution limit’ of 120 km s^{-1} .

The constant term in the fitting function has typically been interpreted as representing the contribution due to interlopers (McKay et al. 2002; Brainerd & Specian 2003). However, we have shown above that $P_{\text{int}}(\Delta V)$ differs strongly from a uniform distribution. The $f_{\text{int}}(L_{\text{host}})$ that follows from integrating the constant term is shown as a dashed line in the panels in the middle column of Fig. 1. As expected, these interloper fractions are systematically too low compared to the true interloper fractions (open circles). Unfortunately, with real data the detailed $P_{\text{int}}(\Delta V)$ is unknown, making it difficult to properly correct for the interlopers. We therefore devised a strategy that aims at tuning the selection criteria to limit the number of interlopers to acceptable levels (cf. SC 3). In order to remove interlopers with large $|\Delta V|$ we still take the constant term into account in the fitting function. As long as the fraction of interlopers is sufficiently small, the remaining interlopers should not strongly effect the kinematics.

In addition to discrete measurements of σ_{sat} for several independent bins in L_{host} , we also use a method that fits $\sigma_{\text{sat}}(L_{\text{host}})$ to all host-satellite pairs simultaneously. We parameterize the relation between σ_{sat} and L_{host} with a quadratic form in the logarithm:

$$\log \sigma_{\text{sat}} = \log \sigma_{10} + a_1 \log L_{10} + a_2 (\log L_{10})^2 \quad (28)$$

Here $L_{10} = L_{\text{host}}/10^{10} h^{-2} L_{\odot}$ and $\sigma_{10} = \sigma_{\text{sat}}(L_{10})$. Let f_{int} denote the fraction of interlopers, and assume that f_{int} is independent of L_{host} and/or ΔV . Then, the probability that a satellite around a host with luminosity L_{host} has a velocity difference with respect to the host of $\Delta V \text{ km s}^{-1}$ is given by

$$P(\Delta V) = \frac{f_{\text{int}}}{2(\Delta V)_s} + \frac{(1 - f_{\text{int}})}{\varpi} \exp \left[-\frac{\Delta V^2}{2\sigma_{\text{eff}}^2} \right] \quad (29)$$

Here $\sigma_{\text{eff}} = \sqrt{\sigma_{\text{sat}}^2 + 120^2}$ defines the *effective* dispersion that takes account of the velocity errors, and

$$\varpi = \sqrt{2\pi} \sigma_{\text{eff}} \operatorname{erf} \left[\frac{(\Delta V)_s}{\sqrt{2} \sigma_{\text{eff}}} \right] \quad (30)$$

so that (29) is properly normalized to unity over the valid range $|\Delta V| \leq (\Delta V)_s$. We use Powell’s directional set method to find the parameters $(\sigma_{10}, a_1, a_2, f_{\text{int}})$ that maximize the likelihood $\mathcal{L} \equiv \sum_i \ln[P(\Delta V_i)]$ where the summation is over all satellites. It is this maximum likelihood method that we use to parameterize the satellite kinemat-

ics in our adaptive selection criterion (SC 3) introduced in Section 3.1. Columns (8)–(11) of Table 2 list the best-fit values for f_{int} (to be compared to the true interloper fraction listed in column 7), σ_{10} , a_1 and a_2 obtained from fitting the host-satellite pairs extracted from the MGRS. The dashed lines in the lower panels of Fig. 4 indicate the corresponding best-fit $\sigma_{\text{sat}}(L_{\text{host}})$.

With SC 1 the satellite kinematics can only be measured accurately over about a factor 5 in luminosity (cf. Brainerd & Specian 2003). Therefore, when fitting $\sigma_{\text{sat}}(L_{\text{host}})$ using the maximum likelihood method we keep a_2 fixed at zero, such that eq. (28) reduces to a simple power-law. As can be seen, the maximum likelihood method and the discrete Gaussian-plus-constant fits yield satellite velocity dispersions in good agreement with each other. Because f_h and f_s are not equal to unity with SC 1, it is difficult to compute expectation values for $\sigma_{\text{sat}}(L_{\text{host}})$ based on the CLF. In order to assess the impact of interlopers and non-central hosts we therefore compute the velocity dispersion of the true satellite galaxies directly from the MGRS: open circles in the upper-left panel of Fig. 4 correspond to $\sqrt{\frac{1}{N} \sum_{i=1}^N (\Delta V_i)^2 - 120^2}$ where the summation is over all true satellites (excluding interlopers and satellites around non-central hosts). The best-fit $\sigma_{\text{sat}}(L)$, both from the maximum likelihood method as well as from the discrete Gaussian-plus-constant fits, are in good agreement with these true values, indicating that the incomplete correction for interlopers does not significantly influence the satellite kinematics.

In the case of SC 2, the luminosity range over which accurate measurements of $\sigma_{\text{sat}}(L)$ can be obtained has increased to almost one and a half orders of magnitude. The $\sigma_{\text{sat}}(L)$ is in reasonable agreement with the expectation values computed using eq. (21) with $R_p = r_{\text{vir}}$ (thick solid line), except for the lowest luminosity bins. This is due to the large fraction of (excess) interlopers and to the presence of satellite galaxies around non-central hosts. Especially the latter can cause a significant overestimate of the true $\sigma_{\text{sat}}(L)$. Nevertheless, despite an interloper-fraction of 39 percent, SC 2 allows one to recover the expected $\sigma_{\text{sat}}(L)$ with reasonable accuracy.

The results for SC 3 are even more promising. Because of the large number of faint host galaxies, $\sigma_{\text{sat}}(L_{\text{host}})$ can be measured over two and a half orders of magnitude down to $\sim 4 \times 10^8 h^{-2} L_{\odot}$. The σ_{sat} obtained are in good agreement with the expectation values (solid line, computed using eq. (21) with $R_p = 0.375 r_{\text{vir}}$), even in the regime where $\sigma_{\text{sat}} < 120 \text{ km s}^{-1}$. The dashed line indicates the best-fit $\sigma_{\text{sat}}(L_{\text{host}})$ obtained from the maximum-likelihood method, the corresponding parameters of which are listed in Table 2.

Fitting eq. (28) directly to the expected $\sigma_{\text{sat}}(L_{\text{host}})$ (solid line in the lower right-hand panel of Fig. 4) yields as best-fit parameters: $\sigma_{10} = 178 \text{ km s}^{-1}$, $a_1 = 0.58$, and $a_2 = 0.10$. Thus, the expectation is somewhat steeper than the actual best-fit relation (for which $a_1 = 0.52$). In order to address the significance of this difference we proceed as follows: we construct Monte-Carlo samples based on $\sigma_{10} = 176 \text{ km s}^{-1}$, $a_1 = 0.52$, $a_2 = 0.11$, and $f_{\text{int}} = 0.12$, corresponding to the best-fit values obtained for the MGRS. For each of the 16841 satellite galaxies in the MGRS we randomly draw a ΔV from equation (29), using f_{int} and $\sigma_{\text{sat}}(L_{\text{host}})$ given by eq. (28). To this ΔV we add a Gaus-

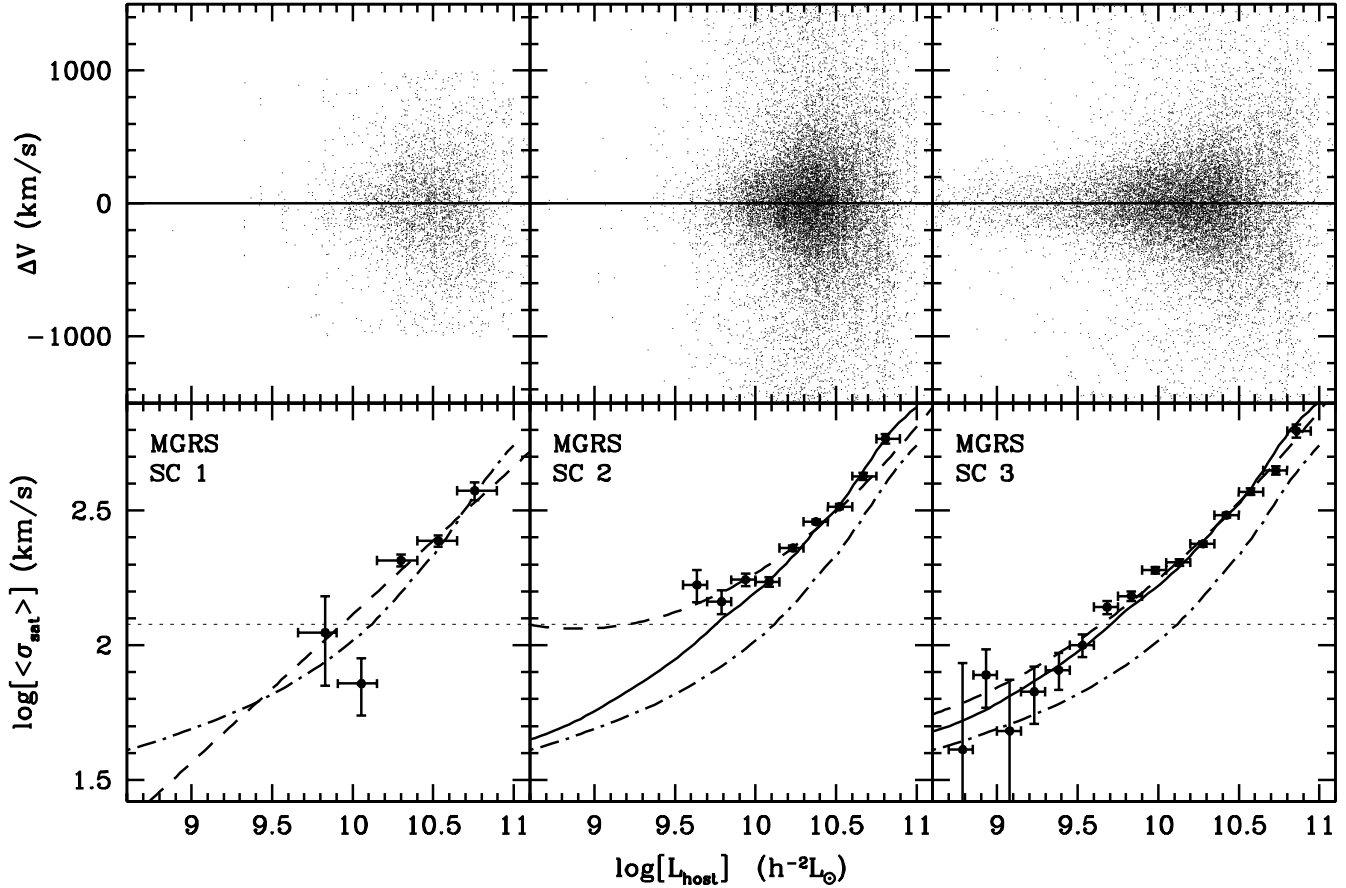


Figure 4. Kinematics of satellite galaxies in the MGRS. Upper panels show scatter plots of ΔV versus L_{host} for the three different host-satellite samples discussed in the text. Lower panels plot the corresponding satellite velocity dispersions. Solid dots with errorbars indicate the $\sigma_{\text{sat}}(L_{\text{host}})$ obtained by fitting $P(\Delta V)$ with the sum of a Gaussian and a constant term. Vertical errorbars are obtained from the covariance matrix of the Levenberg-Marquardt method used for the fitting, and horizontal errorbars indicate the range of host luminosity used. The dashed lines indicate the best-fit $\sigma_{\text{sat}}(L_{\text{host}})$ obtained using the maximum-likelihood method. Open circles in the lower left-hand panel, and solid lines in the lower middle and right-hand panels, indicate the expectation values. Any difference between the $\sigma_{\text{sat}}(L_{\text{host}})$ and these expectation values is due to interlopers, satellites around non-central hosts, and shot noise (see text for detailed discussion). The dot-dashed line indicates the host-averaged mean, and is shown for comparison. The dotted, horizontal line indicates $\sigma_{\text{sat}} = 120 \text{ km s}^{-1}$ which corresponds to the ‘resolution limit’ (due to the errors on ΔV).

sian deviate with standard deviation of 120 km s^{-1} to mimic the velocity errors. This yields a sample of $(\Delta V, L_{\text{host}})$ with the same distributions of L_{host} and N_{sat} as for the MGRS. Next we apply the maximum likelihood method and find the best-fit values of $(\sigma_{10}, a_1, a_2, f_{\text{int}})$. The distributions of these best-fit values, obtained from 1000 of these Monte-Carlo samples, are shown in Fig. 5. Clearly, the maximum-likelihood method accurately recovers the input values of σ_{10} , a_1 , a_2 , and f_{int} (indicated by vertical and horizontal lines) with 1σ errorbars of 3 km s^{-1} , 0.02, 0.04 and 0.005, respectively (and with the errors on a_1 and a_2 somewhat correlated). Given these random errors, the difference between the dashed and solid lines is marginally significant, reflecting the effect of interlopers and (mainly) satellites around non-central hosts. Nevertheless, the difference is sufficiently small, that we conclude that it is possible to use large, flux limited redshift surveys to obtain accurate estimates of the velocity dispersion of satellite galaxies in haloes that span a wide range of masses. However, keep in mind that these are biased with respect to the host-averaged means.

4 RESULTS FOR THE 2dFGRS

We now focus on real data. We use the final, public data release from the 2dFGRS, restricting ourselves only to galaxies with redshifts $0.01 \leq z \leq 0.15$ in the North Galactic Pole (NGP) and South Galactic Pole (SGP) subsamples with a redshift quality parameter $q \geq 3$. This leaves a grand total of 146735 galaxies with a typical rms redshift error of 85 km s^{-1} (Colless et al. 2001). Absolute magnitudes for galaxies in the 2dFGRS are computed using the K-corrections of Madgwick et al. (2002).

4.1 Satellite Kinematics

Applying our adaptive selection criterion (SC 3) to the 2dFGRS yields 8132 host galaxies and 12569 satellite galaxies. The first thing to notice is that these numbers are quite a bit smaller than for the MGRS. In order to check whether this is consistent with cosmic variance, we constructed four independent MGRSs, using different L_{100} and L_{300} simulation boxes. Applying SC 3 to each of these MGRSs yields

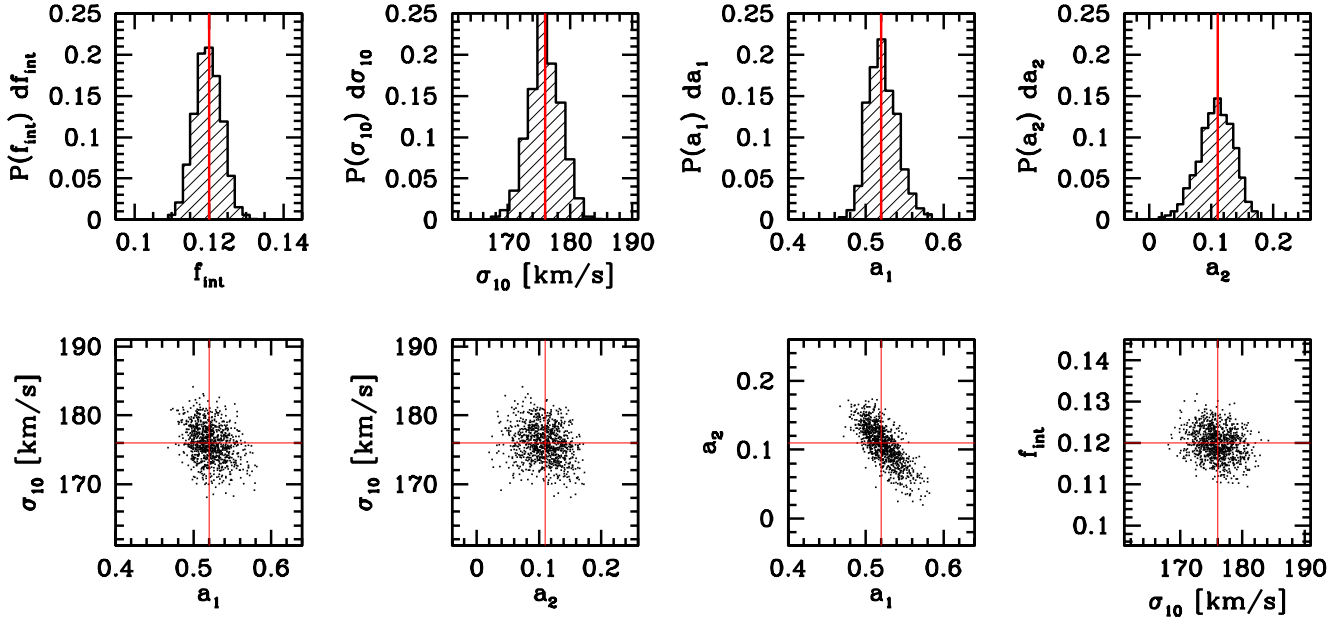


Figure 5. The upper panels plot distributions of f_{int} , σ_{10} , a_1 , and a_2 (as indicated) obtained from 1000 Monte-Carlo simulations. These are used to estimate errors on these parameters obtained from fitting the distribution of $(\Delta V, L_{\text{host}})$ in the MGRS (see text). Thick, vertical lines indicate the input values. The lower panels show two-dimensional scatter plots of these parameters. Horizontal and vertical lines again indicate the input values.

$N_{\text{host}} = 10600 \pm 300$ and $N_{\text{sat}} = 16300 \pm 550$. Clearly, the deficit of host and satellite galaxies in the 2dFGRS compared to our MGRS is very significant. We address the abundances of host and satellite galaxies in much more detail in a forthcoming paper (van den Bosch et al. , in preparation). For the moment, however, we ignore this discrepancy and focus on the kinematics only.

Results are shown in Fig. 6. Solid dots with error bars indicate the $\sigma_{\text{sat}}(L_{\text{host}})$ obtained by fitting a Gaussian-plus-constant. The thick dashed line indicates the best-fit $\sigma_{\text{sat}}(L_{\text{host}})$ obtained using the maximum likelihood method (with the best-fit parameters listed in Table 2). The gray area indicates the expectation values obtained from our CLF where the upper and lower boundaries correspond to ‘random’ and ‘constrained’ luminosity sampling, respectively. Clearly, the satellite kinematics obtained from the 2dFGRS are in excellent agreement with these predictions. For comparison, the solid line indicates the expectation value obtained using host-averaging (see discussion in Section 3.2). As already discussed, the flux limit of the 2dFGRS and the satellite-averaging results in a significant bias of the measured $\sigma_{\text{sat}}(L_{\text{host}})$ compared to this host-averaged mean.

The expectation values (gray area) are computed assuming no spatial bias between satellite galaxies and the dark matter mass distribution. As we have shown in Section 3.2.1, the effect of spatial bias, and the resulting velocity bias, is small compared to the uncertainties resulting from the luminosity sampling. In other words, deviations of the true $n_{\text{sat}}(r)$ from the NFW distribution assumed here will not significantly affect our main conclusion that the satellite kinematics of the 2dFGRS are in excellent agreement with predictions based on our CLF.

5 SUMMARY

Previous attempts to measure the kinematics of satellite galaxies have mainly focussed on isolated spiral galaxies. Using detailed mock galaxy redshift surveys we investigated to what extent a similar analysis can be extended to include a much wider variety of systems, from isolated galaxies to massive groups and clusters. Our method is based on the assumption that satellite galaxies are an isotropic, steady-state tracer population orbiting within spherical NFW dark matter haloes, and that the brightest galaxy in each halo resides at rest at the halo centre. These assumptions are, at least partially, supported by both observations of cluster galaxies (i.e., van der Marel et al. 2000) and by numerical simulations of dark matter subhaloes (see Diemand et al. 2004 and references therein). Although there are most definitely systems in which one or more of these assumptions break down, we hypothesized that with sufficiently large samples of host-satellite pairs the occasional ‘perturbed’, non-relaxed, system will not significantly influence the results. In addition, we have shown that the assumption of orbital isotropy only influences the velocity dispersion of satellite galaxies at the few percent level.

We used the conditional luminosity function (CLF) formalism to make predictions of the observed velocity dispersion of satellite galaxies around host galaxies of different luminosity. We showed that the satellite weighting, which occurs naturally when stacking many host-satellite pairs to increase signal-to-noise, introduces a bias towards higher $\sigma_{\text{sat}}(L_{\text{host}})$ compared to the true, host-averaged mean. A further bias, in the same direction, is introduced when using flux-limited, rather than volume-limited surveys. Finally, we demonstrated that most of the uncertainty in interpreting the measured $\sigma_{\text{sat}}(L_{\text{host}})$ owes to the unknown second moment of the conditional probability distribution $P(M|L_{\text{host}})$:

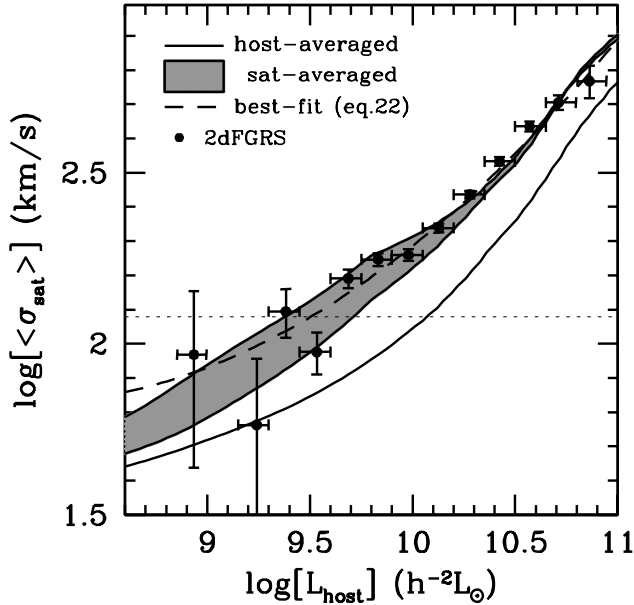


Figure 6. Solid dots with errorbars indicate the best-fit $\sigma_{\text{sat}}(L_{\text{host}})$ obtained from the 2dFGRS data. The dashed line indicates the corresponding best-fit function of the form (28) obtained using the maximum likelihood method described in the text. The gray area indicates the expectation values obtained using the CLF, where the upper and lower boundary correspond to ‘random’ and ‘constrained’ luminosity sampling, respectively. The solid line indicates the corresponding host-averaged mean $\sigma_{\text{sat}}(L_{\text{host}})$ and is shown for comparison.

typically a larger second moment yields higher expectation values. The CLF formalism is the ideal tool to properly take all these various biases into account.

An important, additional problem with the interpretation of satellite kinematics is how to deal with the presence of interlopers (i.e., galaxies selected as satellites but which are not physically associated with the halo of the host galaxy) and non-central hosts (i.e., galaxies that are selected as host galaxies, but which are not the central, brightest galaxy in their own halo). The abundances of interlopers and non-central hosts depend strongly on the selection criteria. Using MGRSs, constructed from the CLF, we investigated different selection criteria for hosts and satellites. The first one is identical to the criteria used by McKay et al. (2002) and Brainerd & Specian (2003) in their analyses of satellite kinematics in the SDSS and 2dFGRS, respectively. Applied to our MGRS it yields only 3876 satellites, which only allows an analysis of satellite kinematics around host galaxies that span a factor 5 in luminosity. Although the fraction of non-central hosts is, with one percent, negligible, the fraction of interlopers is 27 percent. More importantly, the interloper fraction increases strongly with decreasing host-luminosity, reaching values as high as 60 percent around the faintest host galaxies. Contrary to what has been assumed in the past (e.g., McKay et al. 2002; Brainerd & Specian 2003), the velocity distribution of interlopers is strongly peaked towards $\Delta V = 0$, similar to the distribution of true satellite galaxies. This makes a detailed correction for interlopers extremely difficult. In fact, the method used thus far, based on

the assumption of a uniform $P(\Delta V)$, typically underpredicts the true interloper fraction by ~ 50 percent.

We therefore devised an iterative, adaptive selection criterion that yields large numbers of hosts and satellites with as few interlopers and non-central hosts as possible. Applying these criteria to our MGRS yields 16750 satellites around 10483 hosts. In addition, the interloper fraction is only 14 percent and, more importantly, does not vary significantly with host luminosity. Because of the much larger numbers involved, and the reduced fraction of interlopers, satellite kinematics can be analyzed over two and a half orders of magnitude in L_{host} . The resulting $\sigma_{\text{sat}}(L_{\text{host}})$ is found to be in good agreement with the expectation values, indicating that the interlopers and non-central hosts do not significantly distort the measurements.

Finally, we applied our adaptive selection criteria to the 2dFGRS. The resulting satellite kinematics are in excellent agreement with predictions based on the CLF formalism, once the various biases discussed above are taken into account. We therefore conclude that the observed kinematics of satellite galaxies provide virtually independent, dynamical confirmation of the average mass-to-light ratios inferred by Yang et al. (2003a) and van den Bosch et al. (2003a) from a purely statistical method based on the abundances and clustering properties of galaxies in the 2dFGRS.

ACKNOWLEDGEMENTS

We are grateful to Yipeng Jing for providing us the set of numerical simulations used for the construction of our mock galaxy redshift surveys, and to Darren Croton, Ben Moore, Juerg Diemand, Anna Pasquali, Simon White, and the anonymous referee for useful comments and discussions.

REFERENCES

- Bahcall N.A., Lubin L., Dorman V., 1995, *ApJ*, 447, L81
- Bahcall N.A., Cen R., Davé R., Ostriker J.P., Yu Q., 2000, *ApJ*, 541, 1
- Benson A.J., Cole S., Frenk C.S., Baugh C.M., Lacey C.G., 2000, *MNRAS*, 311, 793
- Berlind A.A., Weinberg D.H., 2002, *ApJ*, 575, 587
- Berlind A.A., et. al., 2003, *ApJ*, 593, 1
- Binney J.J., Tremaine S.D., 1987, *Galactic Dynamics* (Princeton: Princeton Univ. Press)
- Brainerd T.G., Specian M.A., 2003, *ApJ*, 593, L7
- Carlberg R.G., Yee H.K.C., Ellingson E., Abraham R., Gravel P., Morris S., Pritchet C.J., 1996, *ApJ*, 462, 32
- Carlberg R.G., et al., 1997a, *ApJ*, 485, L13
- Carlberg R.G., Yee H.K.C., Ellingson E., 1997b, *ApJ*, 478, 462
- Colin P., Klypin A., Kravtsov A.V., Khokhlov M., 1999, *ApJ*, 532, 32
- Colless M., The 2dFGRS team, 2001, *MNRAS*, 328, 1039
- Davis M., Efstathiou G., Frenk C.S., White S.D.M., 1985, *ApJ*, 292, 371
- De Lucia G., Kauffmann G., Springel V., White S.D.M., Lanzoni B., Stoehr F., Tormen G., Yoshida N., 2004, *MNRAS*, 348, 333
- Diemand J., Moore B., Stadel J., 2004, preprint (astro-ph/0402160)
- Eke V.R., Navarro J.F., Steinmetz M., 2001, *ApJ*, 554, 114
- Eke V.R., The 2dFGRS team, 2004, preprint (astro-ph/0402566)

- Erickson L.K., Gottesman S.T., Hunter J.H., 1987, *Nature*, 325, 779
- Evans N.W., Wilkinson M.I., 2000, *MNRAS*, 316, 929
- Evans N.W., Wilkinson M.I., Guhathakurta P., Grebel E.K., Vogt S.S., 2000, *ApJ*, 540, L9
- Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 1998, *MNRAS*, 300, 146
- Gradshteyn I.S., Ryzhik I.M., 1980, *Table of Integrals, Series, and Products*. Academic Press, New York
- Jing Y.P., 2002, *MNRAS*, 335, L89
- Jing Y.P., Suto, Y., 2002, *ApJ*, 574, 538
- Klypin A., Gottlöber S., Kravtsov A.V., Khokhlov A.M., 1999, *ApJ*, 516, 530
- Kravtsov A.V., Berlind A.A., Wechsler R.H., Klypin A.A., Gottlöber S., Allgood B., Primack J.R., 2003, preprint (astro-ph/0308519)
- Lin D.N.C., Jones B.F., Klemola A.R., 1995, *ApJ*, 439, 652
- Lin Y.-T., Mohr J.J., Stanford S.A., 2004, preprint (astro-ph/0402308)
- Little B., Tremaine S., 1987, 320, 493
- Maddox S.J., Efstathiou G., Sutherland W.J., Loveday L., 1990, *MNRAS*, 243, 692
- Madgwick D.S., The 2dFGRS team, 2002, *MNRAS*, 333, 133
- McKay T.A. et al., 2002, *ApJ*, 571, L85
- Moore B., Governato G., Quinn T., Stadel J., Lake G., 1998, *ApJ*, 499, L5
- Navarro J.F., Frenk C.S., White S.D.M., 1997, *ApJ*, 490, 493
- Norberg P., The 2dFGRS team, 2002a, *MNRAS*, 332, 827
- Norberg P., The 2dFGRS team, 2002b, *MNRAS*, 336, 907
- Okamoto T., Habe A., 1999, *ApJ*, 516, 591
- Prada F., et al., 2003, *ApJ*, 598, 260
- Rhines K., Geller M.J., Diaferio A., Kurtz M.J., Jarrett T.H., 2004, preprint (astro-ph/0402242)
- Sanderson A.J.R., Ponman T.J., 2003, *MNRAS*, 345, 1241
- Scoccimarro R., Sheth R.K., Hui L., Jain B., 2001, *ApJ*, 546, 20
- Seljak U., 2000, *MNRAS*, 318, 203
- Sheth R.K., Mo H.J., Tormen G., 2001a, *MNRAS*, 323, 1
- Sheth R.K., Tormen, G., 2002, *MNRAS*, 329, 61
- Springel V., White S.D.M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Vale A., Ostriker J.P., 2004, preprint (astro-ph/0402500)
- van den Bosch F.C., Yang X., Mo H.J., 2003a, *MNRAS*, 340, 771
- van den Bosch F.C., Mo H.J., Yang X., 2003b, *MNRAS*, 345, 923
- van der Marel R.P., Magorrian J., Carlberg R.G., Yee H.K.C., Ellingson E., 2000, *AJ*, 119, 2038
- Yang X., Mo H.J., van den Bosch F.C., 2003a, *MNRAS*, 339, 1057
- Yang X., Mo H.J., Jing Y.P., van den Bosch F.C., Chu Y., 2003b, *MNRAS*, in press (astro-ph/0303524)
- York D., et al., 2000, *AJ*, 120, 1579
- Zaritsky D., Smith R., Frenk C.S., White S.D.M., 1993, *ApJ*, 405, 464
- Zaritsky D., White S.D.M., 1994, *ApJ*, 435, 599
- Zaritsky D., Smith R., Frenk C.S., White S.D.M., 1997, *ApJ*, 478, 39
- Zwicky F., 1933, *Helv. Phys. Acta*, 6, 110
- Zwicky F., 1937, *ApJ*, 86, 217

APPENDIX A: THE CONDITIONAL LUMINOSITY FUNCTION

The construction of our mock galaxy redshift surveys uses the conditional luminosity function (CLF) to indicate how many galaxies of given luminosity occupy a halo of given mass. The CLF formalism was introduced by Yang et al. (2003a) and van den Bosch et al. (2003a), and we refer the reader to these papers for more details. Here we briefly

summarize the main ingredients and we present the parameterization used.

The CLF is parameterized by a Schechter function:

$$\Phi(L|M)dL = \frac{\tilde{\Phi}^*}{\tilde{L}^*} \left(\frac{L}{\tilde{L}^*} \right)^{\tilde{\alpha}} \exp(-L/\tilde{L}^*) dL, \quad (\text{A1})$$

where $\tilde{L}^* = \tilde{L}^*(M)$, $\tilde{\alpha} = \tilde{\alpha}(M)$ and $\tilde{\Phi}^* = \tilde{\Phi}^*(M)$ are all functions of halo mass M . We write the average, total mass-to-light ratio of a halo of mass M as

$$\left\langle \frac{M}{L} \right\rangle (M) = \frac{1}{2} \left(\frac{M}{L} \right)_0 \left[\left(\frac{M}{M_1} \right)^{-\gamma_1} + \left(\frac{M}{M_1} \right)^{\gamma_2} \right], \quad (\text{A2})$$

This parameterization has four free parameters: a characteristic mass M_1 , for which the mass-to-light ratio is equal to $(M/L)_0$, and two slopes, γ_1 and γ_2 , that specify the behavior of $\langle M/L \rangle$ at the low and high mass ends, respectively. Motivated by observations (Bahcall, Lubin & Norman 1995; Bahcall et al. 2000; Sanderson & Ponman 2003; Eke et al. 2004), which indicate a flattening of $\langle M/L \rangle(M)$ on the scale of galaxy clusters, we set $\langle M/L \rangle(M) = (M/L)_{\text{cl}}$ for haloes with $M \geq 10^{14} h^{-1} M_\odot$. With $(M/L)_{\text{cl}}$ specified, the value for γ_2 derives from requiring continuity in $\langle M/L \rangle(M)$ across $M = 10^{14} h^{-1} M_\odot$.

A similar parameterization is used for the characteristic luminosity $\tilde{L}^*(M)$:

$$\frac{M}{\tilde{L}^*(M)} = \frac{1}{2} \left(\frac{M}{L} \right)_0 f(\tilde{\alpha}) \left[\left(\frac{M}{M_1} \right)^{-\gamma_1} + \left(\frac{M}{M_2} \right)^{\gamma_3} \right], \quad (\text{A3})$$

with

$$f(\tilde{\alpha}) = \frac{\Gamma(\tilde{\alpha} + 2)}{\Gamma(\tilde{\alpha} + 1, 1)}. \quad (\text{A4})$$

Here $\Gamma(x)$ is the Gamma function and $\Gamma(a, x)$ the incomplete Gamma function. This parameterization has two additional free parameters: a characteristic mass M_2 and a power-law slope γ_3 . For $\tilde{\alpha}(M)$ we adopt a simple linear function of $\log(M)$,

$$\tilde{\alpha}(M) = \alpha_{15} + \eta \log(M_{15}), \quad (\text{A5})$$

with M_{15} the halo mass in units of $10^{15} h^{-1} M_\odot$, $\alpha_{15} = \tilde{\alpha}(M_{15} = 1)$, and η describes the change of the faint-end slope $\tilde{\alpha}$ with halo mass. Finally, we introduce the mass scale M_{min} below which we set the CLF to zero; i.e., we assume that no stars form inside haloes with $M < M_{\text{min}}$. Motivated by reionization considerations (see Yang et al. 2003a for details) we adopt $M_{\text{min}} = 10^9 h^{-1} M_\odot$ throughout.

In this paper we use a CLF with the following parameters: $M_1 = 10^{11.12} h^{-1} M_\odot$, $M_2 = 10^{11.71} h^{-1} M_\odot$, $(M/L)_0 = 85h$ (M/L) $_\odot$, $\gamma_1 = 1.55$, $\gamma_2 = 0.46$, $\gamma_3 = 0.69$, $\eta = -0.29$ and $\alpha_{15} = -0.99$. This model is different from those listed in van den Bosch et al. (2003a) as it yields a higher, average mass-to-light ratio on the scale of clusters. As shown in Yang et al. (2003b), this is in better agreement with the observed pairwise peculiar velocity dispersions of 2dFGRS galaxies (see also van den Bosch, Mo & Yang 2003b). However, none of the results presented in this paper are sensitive to our choice of the CLF. We verified that MGRSSs based on either of the CLFs listed in van den Bosch et al. (2003a) yield virtually identical results to those presented here.

APPENDIX B: PROBABILITY DISTRIBUTION OF LUMINOSITY OF BRIGHTEST GALAXY IN A DARK MATTER HALO.

Define L_c as the luminosity of the *brightest* galaxy in a halo of mass M . It is convenient to write the conditional probability distribution $P(L_c|M)dL_c$ in terms of the conditional luminosity function $\Phi(L_c|M)dL_c$ and a new function $f(L_c, M)$ which depends on how galaxy luminosities are drawn from the CLF:

$$P(L_c|M)dL_c = \Phi(L_c|M) f(L_c, M) dL_c \quad (\text{B1})$$

In the case of ‘constrained’ drawing, L_c has an expectation value given by eq. (5), and it is straightforward to show that

$$f(L_c, M) = \begin{cases} 1 & \text{if } L_c \geq L_1(M) \\ 0 & \text{if } L_c < L_1(M) \end{cases} \quad (\text{B2})$$

with $L_1(M)$ as defined by eq. (4). Note that in the MGRS used in this paper the luminosity of the brightest galaxy in each halo is always drawn constrained. Therefore, when computing expectation values for $\sigma_{\text{sat}}(L_{\text{host}})$ in the MGRS, we use $P(L_c|M)$ with $f(L_c, M)$ given by eq. (B2).

In the case of ‘random’ drawing, the situation is more complicated. The probability that a galaxy drawn at random from the CLF has a luminosity less than L_c is given by

$$P(< L_c|M) = 1 - \frac{1}{\langle N \rangle_M} \int_{L_c}^{\infty} \Phi(L|M)dL \quad (\text{B3})$$

with $\langle N \rangle_M$ the mean number of galaxies in a halo of mass M given by eq. (1). In a halo with N galaxies, the probability that the brightest galaxy has $L < L_c$ is simply $[P(< L_c|M)]^N$. Differentiating with respect to L_c yields the probability $P_N(L_c|M)dL_c$ that after N drawings the brightest galaxy has a luminosity in the range $L_c \pm L_c/2$. The full probability $P(L_c|M)dL_c$ follows from summing $P_N(L_c|M)dL_c$ over N , properly weighted by the probability $P(N|M)$ that a halo of mass M contains N galaxies. This yields

$$f(L_c, M) = \frac{1}{\langle N \rangle_M} \sum_{N=1}^{\infty} N P(N|M) [P(< L_c|M)]^{N-1} \quad (\text{B4})$$

If $\langle N \rangle_M \leq 1$ then

$$P(N|M) = \begin{cases} 1 - \langle N \rangle_M & \text{if } N = 0 \\ \langle N \rangle_M & \text{if } N = 1 \end{cases} \quad (\text{B5})$$

(see Section 2.2), so that $f(L_c, M) = 1$. For $\langle N \rangle_M > 1$ we have that

$$P(N|M) = \frac{(\langle N \rangle_M - 1)^{N-1}}{(N-1)!} \exp(1 - \langle N \rangle_M) \quad (\text{B6})$$

(see Section 2.2). Substituting (B6) in (B4) and using that

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} (k+1) = (1+x) \exp(x) \quad (\text{B7})$$

(see eq. [1.212] in Gradshteyn & Ryzhik, 1980) yields

$$f(L_c, M) = \left(1 - \frac{\zeta}{\langle N \rangle_M} \right) \exp(-\zeta) \quad (\text{B8})$$

with

$$\zeta = \frac{\langle N \rangle_M - 1}{\langle N \rangle_M} \int_{L_c}^{\infty} \Phi(L|M)dL \quad (\text{B9})$$